

CHISEL

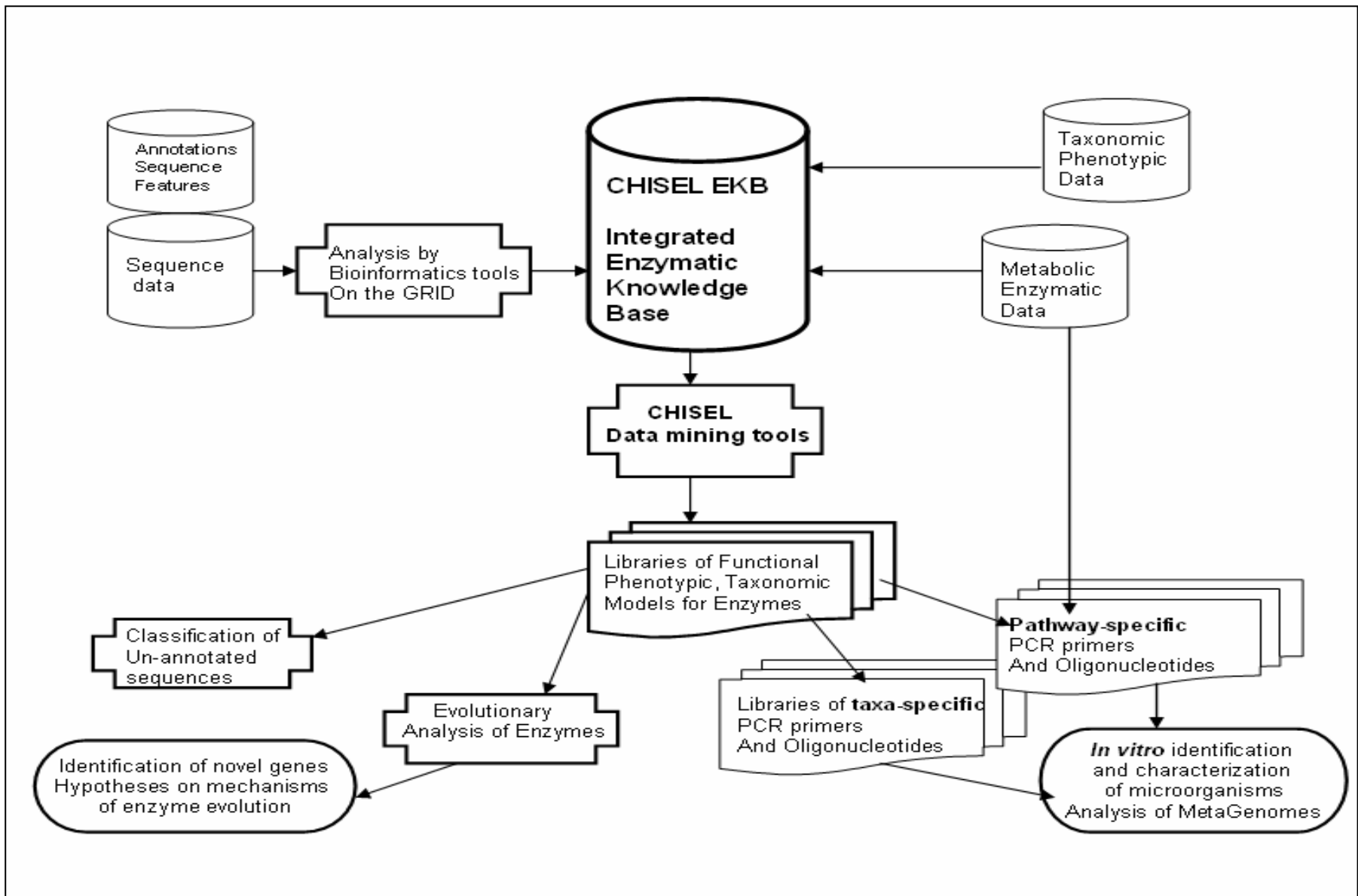
MCS

Argonne National Laboratory

Motivation

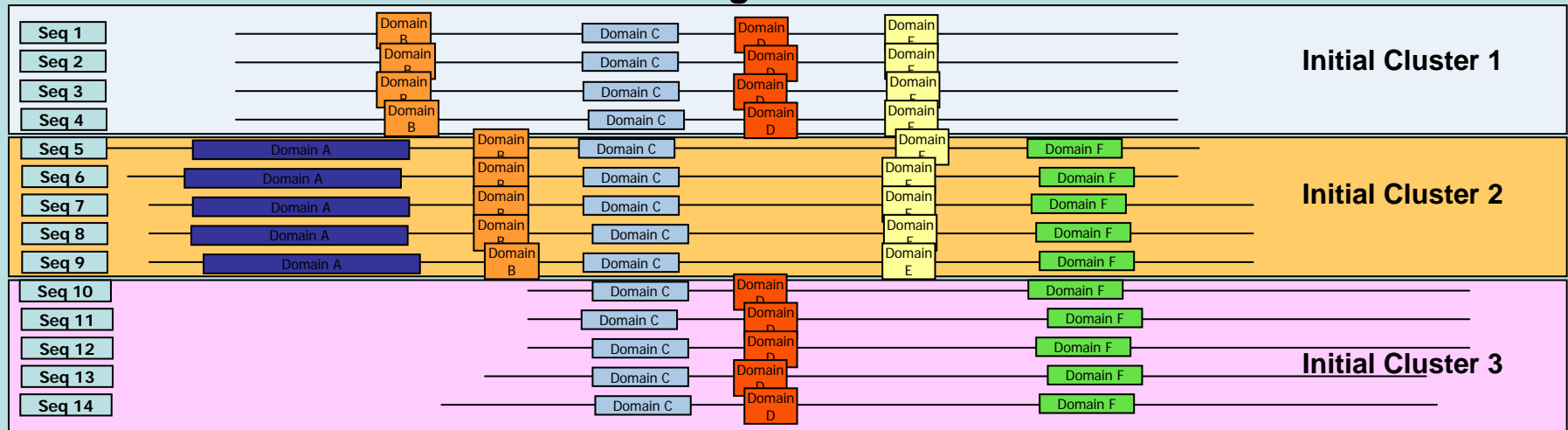
Comparative analysis of wide spectrum of organisms is essential for understanding their adaptation to the environment. All the organisms, although they share a common ancestor adapt themselves to the environment changes where they are living and evolve with different variations of function, organization and phenotypic features. Recent progress in genomics, bioinformatics and physiological studies now allows for systematic exploration of the adaptive mechanisms that lead to diversification of biological systems. Therefore, the development of the scientific framework for studying the evolution processes and phenotypic variation is essential for the interpretation of such events.

CHISEL Workflow



CHISEL Algorithm: Level 1

First Level: Domain Alignment

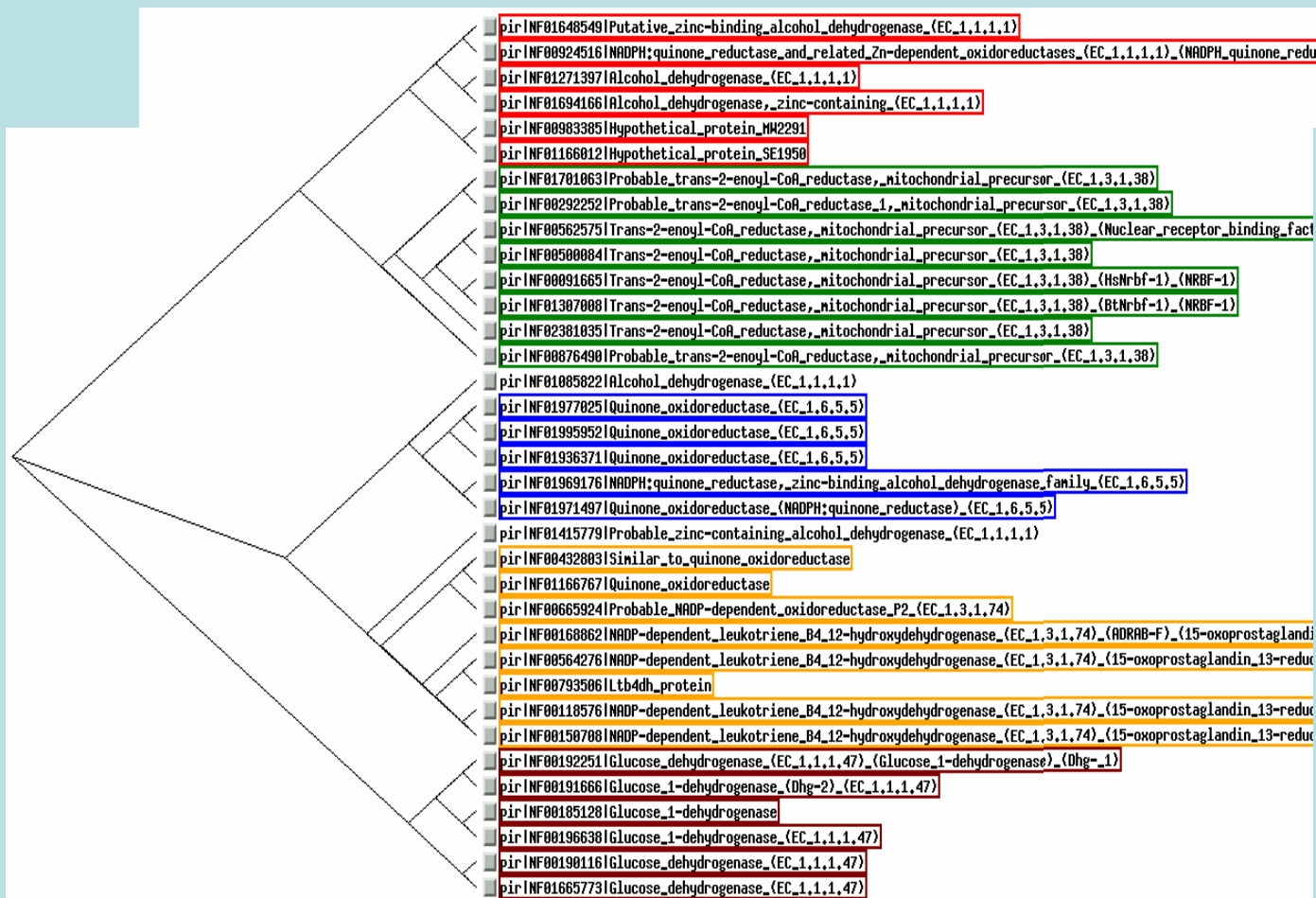


At this level, it is important to differentiate and cluster sequences having varying functional domain composition and sequence lengths. These differences may be due to evolution of the organisms.

CHISEL Algorithm: Level 2 and 3

Second Level: Functional Classification

Third Level: Taxonomy Classification



Alcohol dehydrogenase superfamily with similar domain composition. Members of the initial cluster perform a variety of clusters. At this level, the initial cluster is broken into sub-clusters which will be further refined based on function and taxonomy.

CHISEL Homepage

CHISEL

high-resolution evolutionary analysis of enzymatic sequences

[Home](#)[Search](#)[Classify](#)[Contact us](#)[Download](#)[Help](#)

Tools

[PhyloBlocks](#)[BloBla](#)

Related Resources

[PUMA2](#)[GNARE](#)[Reference & Supplemental
Material](#)

CHISEL is an integrated framework for identification and characterization of taxonomic and phenotypic variations of enzymes.

The taxonomy-specific clusters of enzymatic sequences are represented as computational models (HMM profiles, PSSMs, consensus sequences and multiple sequence alignments) available for [download](#).

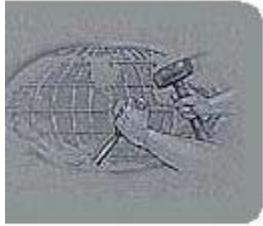
Currently Chisel contains **8575** models that are both taxonomy and function specific. You can search Chisel clusters based on taxonomy, phenotypic characteristics of the organisms, metabolic pathways and sequence feature characteristics.

The [classification](#) page allows users to submit a sequence for prediction of function and possible taxonomy of the organism of origin.

Chisel was successfully used for the needs of evolutionary analysis, biodefense research and analyses of metagenomes. [View](#)

[Disclaimer](#)[Security Notice](#)

Three Searching options



CHISEL

high-resolution evolutionary analysis of enzymatic sequences

[Home](#)[Search](#)[Classify](#)[Contact us](#)[Download](#)[Help](#)

There are different type of searches you can do:

Search based on Sequence ID and Protein Family



Search based on:

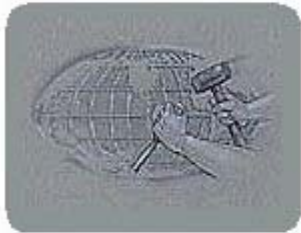
- Phenotypic Characteristics (niche, salinity, temperature, pathogenicity, and more)
- Tax Lineage
- Participation in Metabolic Pathways



Project taxonomy specific clusters onto metabolic pathways



CHISEL Search based on Seq. ID and Protein Family



CHISEL

high-resolution evolutionary analysis of enzymatic sequences

[Home](#)[Search](#)[Classify](#)[Contact us](#)[Download](#)[Help](#)

Choose one of the following families below from which the analysis has to be done

- ☒ **Pir**
- ☐ **Cog** (Not implemented yet)
- ☐ **Tigrfam** (Not implemented yet)

Select one of the fields below for retrieving its related superfamilies from pir

EC Number ▼

1.1.1.2

retrieve

reset

INPUT EXAMPLES

EC Number	1.1.1.2
NCBI gi Number	49480910
PIR Superfamily ID	SF000193

Get sequences based on EC Number, PIR SF or GI for Chisel Analysis and cluster development.

Chisel Search by seq or protein family

Report for the SF000193 family of sequences

PIR Superfamily	SF000193	General Seed Cluster Information
PIR Superfamily name	pyrroline-5-carboxylate reductase	
Number of Sequences	285	
EC distribution	EC: 1.5.1.2	
Interpro Domains	IPR004455 NADP oxidoreductase, coenzyme F420-dependent	(279 proteins)
	IPR000583 Glutamine amidotransferase, class-II	(7 proteins)
	IPR006115 6-phosphogluconate dehydrogenase, NAD-binding	(2 proteins)
	IPR008948 L-Aspartase-like	(11 proteins)
	IPR011128 NAD-dependent glycerol-3-phosphate dehydrogenase, N-terminal	(1 proteins)
	IPR000304 Delta 1-pyrroline-5-carboxylate reductase	(280 proteins)
	IPR001865 Ribosomal protein S2	(1 proteins)
	IPR002197 Helix-turn-helix, Fis-type	(4 proteins)
	IPR006035 Arginase/agmatinase/formiminoglutamase	
	IPR000719 Protein kinase	
	IPR000594 UBA/THIF-type NAD/FAD binding fold	
	IPR000205 NAD-binding site	
	IPR002844 Methylene-5,6,7,8-tetrahydromethanopterin dehydrogenase	(3 proteins)
	IPR002086 Aldehyde dehydrogenase	(2 proteins)
	IPR000870 Homoserine kinase	(7 proteins)
Taxonomy distribution	A (15 proteins) B (204 proteins) E (66 proteins)	

Interpro Domain
Composition of
Sequences

Sequence List to
submit to Cluster
Analysis

Send to Chisel

Send to Phylo Blocks

☐ Select the sequences for performing chisel analysis (230 sequences are chosen for analysis)

☒ [gi 51974827](#)

EC 1.5.1.2

B

Bacillus cereus

pyrroline-5-carboxylate reductase [Bacillus cereus E33L]

Chisel Search by seq or protein family

Clusters of proteins on which blocks can be created

Taxonomic and function specific
Chisel clusters generated giving
similar characteristics of the
sequences within the clusters

Cluster 1

Evidence: *Common Interpro domains, taxonomy and function composition*

gi 71852692	(256)	EC_1.5.1.2	B	Streptococcus pyogenes	Streptococcus pyogenes	pyrroline-5-carboxylate reductase [Streptococcus pyogenes MGAS5005]	IPR000304 IPR004455
gi 19745290	(256)	EC_1.5.1.2	B	Streptococcus pyogenes	Streptococcus pyogenes	putative pyrroline carboxylate reductase [Streptococcus pyogenes MGAS8232]	IPR000304 IPR004455
gi 28810246	(256)	EC_1.5.1.2	B	Streptococcus pyogenes	Streptococcus pyogenes	putative pyrroline carboxylate reductase [Streptococcus pyogenes SSI-1]	IPR000304 IPR004455
gi 50902564	(256)	EC_1.5.1.2	B	Streptococcus pyogenes	Streptococcus pyogenes	Pyrroline-5-carboxylate reductase [Streptococcus pyogenes MGAS10394]	IPR000304 IPR004455

Cluster 2

Evidence: *Common Interpro domains, taxonomy and function composition*

gi 11037704	(290)	EC_1.5.1.2	B	Rhizobiales	Bradyrhizobium japonicum	delta 1-pyrroline-5-carboxylate reductase [Bradyrhizobium japonicum]	IPR000304 IPR004455
gi 27355737	(292)	EC_1.5.1.2	B	Rhizobiales	Bradyrhizobium japonicum	delta 1-pyrroline-5-carboxylate reductase [Bradyrhizobium japonicum USDA 110]	IPR000304 IPR004455
gi 39651285	(280)	EC_1.5.1.2	B	Rhizobiales	Rhodopseudomonas palustris	pyrroline-5-carboxylate reductase [Rhodopseudomonas palustris CGA009]	IPR000304 IPR004455
gi 13475432	(271)	EC_1.5.1.2	B	Rhizobiales	Mesorhizobium loti	pyrroline-5-carboxylate reductase [Mesorhizobium loti MAFF303099]	IPR000304 IPR004455

Cluster 3

Evidence: *Common Interpro domains, taxonomy and function composition*

gi 15595590	(273)	EC_1.5.1.2	B	Gammaproteobacteria	Pseudomonas aeruginosa	pyrroline-5-carboxylate reductase [Pseudomonas aeruginosa PAO1]	IPR000304 IPR002197 IPR004455
gi 24986881	(272)	EC_1.5.1.2	B	Gammaproteobacteria	Pseudomonas putida	pyrroline-5-carboxylate reductase [Pseudomonas putida KT2440]	IPR000304 IPR002197 IPR004455
gi 28855415	(272)	EC_1.5.1.2	B	Gammaproteobacteria	Pseudomonas syringae group genomosp. 3	pyrroline-5-carboxylate reductase [Pseudomonas syringae pv. tomato str. DC3000]	IPR000304 IPR002197 IPR004455
gi 133757117	(272)	EC_1.5.1.2	B	Gammaproteobacteria	Methylobacillus	pyrroline-5-carboxylate reductase [Methylobacillus	IPR000304 IPR002197 IPR004455

Chisel Search by Phenotypes or metabolic pathways

Select the parameters below to refine your search of Chisel clusters containing:

You Must Select a Kingdom:

1. <input checked="" type="checkbox"/> Bacteria	<input checked="" type="checkbox"/> Archaeae	<input checked="" type="checkbox"/> Eukaryote	2. SuperFamily ID: <input type="text"/>
3. Metabolic Pathway Involvement:	<input type="text" value="NO SELECTION"/>		
4. Search only Taxonomic Specific Clusters:	<input type="radio"/> Yes <input checked="" type="radio"/> No		
5. Tax Lineage Keyword:	<input type="text"/>		
7. EC Number:	<input type="text"/>		
9. Sort By:	<input type="radio"/> Tax ID <input checked="" type="radio"/> Org Name		

6. Select Phenotype Characteristics:	
<input type="radio"/> AND <input checked="" type="radio"/> OR	
a. Environment Niche:	b. Salinity:
<div><div>Aquatic</div><div>Host-associated</div></div>	<div><div>Extreme halophilic</div><div>Mesophilic</div></div>
<input type="checkbox"/> Clusters containing only specified options	<input type="checkbox"/> Clusters containing only specified options
c. Oxygen:	d. Temperature:
<div><div>Aerobic</div><div>Anaerobic</div></div>	<div><div>Hyperthermophilic</div><div>Mesophilic</div></div>
<input type="checkbox"/> Clusters containing only specified options	<input type="checkbox"/> Clusters containing only specified options
e. Pathogenicity:	f. Energy Source:
<div><div>Animal</div><div>Animal, Birds</div></div>	<div><div>chemoautotrophic</div><div>chemoautotrophic or chemoheterotrophic</div></div>
<input type="checkbox"/> Clusters containing only specified options	<input type="checkbox"/> Clusters containing only specified options

Clear

Chisel allows to search specific clusters based on several phenotypic and function characteristic or metabolic pathway involvement

Chisel Search by Phenotypes and Pathways

Chisel displays Cluster characteristics such as EC, Metabolic Pathway, common taxonomy, phenotypic characteristics and more.

Chisel Cluster: EC 1.1.1.1 - Bacteria - *Bacteria* Type [[SF000091](#) 8 B *Bacteria*1] [View Features](#)

Enzyme Classification: [1.1.1.1](#) (alcohol dehydrogenase)
PIR Superfamily: alcohol dehydrogenase [[SF000091](#)]
Pathways Involved: [Glycerophospholipid metabolism](#) [Tyrosine metabolism](#) [Bile acid biosynthesis](#) [1-Gluconeogenesis](#) [Glycerolipid metabolism](#) [Fatty acid metabolism](#) [Aminophosphonate metabolism](#)
Common Taxonomy: *Bacteria*
Reactions: [R00623](#) [R00624](#) [R00754](#) [R01036](#) [R01041](#) [R04805](#) [R04880](#) [R06917](#) [R06927](#) [1.1.1.10](#)
Diseases: [MIM: 103720](#) Alcohol dehydrogenase IB (class I), beta polypeptide 1.1.1.1

Organisms Phenotypes

GI <input type="checkbox"/> All <input type="checkbox"/> Consensus	Sequence Description	Organism	Habitat	Salinity	Oxygen	Temperature	Pathogenicity	Carbon/Energy Source
<input type="checkbox"/> 26453921	Alcohol dehydrogenase	Mycoplasma penetrans	Host-associated	Non-halophilic	Facultative	Mesophilic	Human	chemoorganotrophic
<input type="checkbox"/> 28272178	Alcohol dehydrogenase	Lactobacillus plantarum	Host-associated	---	Facultative	Mesophilic	No	chemoorganotrophic
<input type="checkbox"/> 15459343	Alcohol dehydrogenase	Streptococcus pneumoniae	Multiple	---	Facultative	Mesophilic	Human	chemoorganotrophic
<input type="checkbox"/> 28271191	Alcohol dehydrogenase	Lactobacillus plantarum	Host-associated	---	Facultative	Mesophilic	No	chemoorganotrophic
<input type="checkbox"/> 13475922	Alcohol dehydrogenase	Mesorhizobium loti	Multiple	---	Aerobic	Mesophilic	No	---
<input type="checkbox"/> 14523738	Probable alcohol	Sinorhizobium meliloti	Multiple	---	Aerobic	Mesophilic	No	chemoorganotrophic
<input type="checkbox"/> 141900	Alcohol dehydrogenase	Cupriavidus necator	Terrestrial	---	Aerobic	Mesophilic	No	chemoheterotrophic

Chisel Cluster: EC 1.1.1.1 - Eukaryota - *Drosophila* Type [[SF000092](#) 12 E *Drosophila*1] [View Features](#)

Enzyme Classification: [1.1.1.1](#) (alcohol dehydrogenase)
PIR Superfamily: short-chain dehydrogenase [[SF000092](#)]
Pathways Involved: [Glycerophospholipid metabolism](#) [Tyrosine metabolism](#) [Bile acid biosynthesis](#) [1- and 2-Methylglutamate metabolism](#) [Gluconeogenesis](#) [Glycerolipid metabolism](#) [Fatty acid metabolism](#) [Aminophosphonate metabolism](#)
Common Taxonomy: Eukaryota, Fungi/Metazoa group, Metazoa, Eumetazoa, Bilateria, Coelomata, Protostomia, Pancrustacea, Hexapoda, Insecta, Dicondylia, Pterygota, Neoptera, Endopterygota, Diptera, Ephydroptera, Acalyptratae, Ephydroidea, Drosophilidae, Drosophilinae, Drosophilini, Drosophilina, Drosophila
Reactions: [R00623](#) [R00624](#) [R00754](#) [R01036](#) [R01041](#) [R04805](#) [R04880](#) [R06917](#) [R06927](#) [1.1.1.10](#)
Diseases: [MIM: 103720](#) Alcohol dehydrogenase IB (class I), beta polypeptide 1.1.1.1

Color bar indicates sequences belong to different PIR superfamily cluster.

Organisms Phenotypes

GI <input type="checkbox"/> All <input type="checkbox"/> Consensus	Sequence Description	Organism	Habitat	Salinity	Oxygen	Temperature	Pathogenicity	Carbon/Energy Source
<input type="checkbox"/> 158042	Alcohol dehydrogenase	Drosophila persimilis	---	---	---	---	---	---
<input type="checkbox"/> 8870	Alcohol dehydrogenase	Drosophila pseudoobscura	---	---	---	---	---	---
<input type="checkbox"/> 156841	Alcohol dehydrogenase	Drosophila subobscura	---	---	---	---	---	---
<input type="checkbox"/> 207074	Alcohol dehydrogenase	Drosophila melanogaster	---	---	---	---	---	---

Chisel Search Feature results

[Send to Phylblocks](#)
[Send to ClustalW](#)
[Send to POAVIZ](#)
[View DNA FASTA contig](#)

Chisel Cluster: EC 1.5.1.2 - Bacteria - *Proteobacteria* Type [[SF000193_5_B_Proteobacteria1](#)] %similarity: 39.28]

Enzyme Classification: [1.5.1.2](#) (pyrroline-5-carboxylate reductase)

PIR Superfamily: pyrroline-5-carboxylate reductase [[SF000193](#)]

Pathways Involved: [Urea cycle and metabolism of amino groups](#) [Aminoacyl-tRNA synthetases](#)
[Arginine and proline metabolism](#)








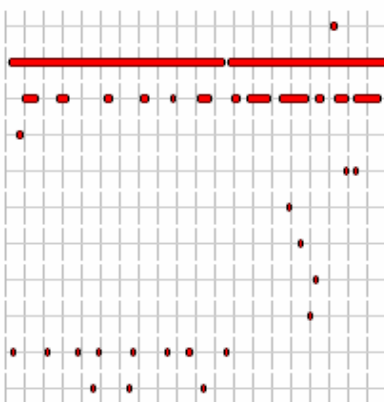







Common Taxonomy: Bacteria, *Proteobacteria*

Reactions: [R01248](#) [R01251](#) [R03291](#) [R03293](#)

Diseases: ---

Chisel also displays Cluster characteristics regarding Feature and structure composition.

Structural and Feature Data

Sequence <input type="checkbox"/> All <input type="checkbox"/> Consensus	GO	SCOP	PDB	Features	CATH
<input type="checkbox"/> GI:7225275 Pyrroline-5-carboxylate reductase Neisseria meningitidis Cofactor-> No Data	GO:0004735 (IEA) GO:0016491 (IEA)	---	PDB XREF: 1YQG Sequence length (263 aa)  2AG8  2AMF  2AHR  1SO8  1U7T 	Sequence length (263 aa)  ACTIVE CENTER DOMAIN HELI MOTIF SALT BRIDGE 1 SALT BRIDGE 2 SALT BRIDGE 3 SALT BRIDGE 4 SALT BRIDGE 5 STRAND TURN 	---
<input type="checkbox"/> GI:59719114 Putative pyrroline-5-carboxylate reductase Neisseria gonorrhoeae Cofactor-> No Data	GO:0004735 (IEA) GO:0016491 (IEA)	---	PDB XREF: ---- Sequence length (263 aa)  2AG8  2AMF  2AHR  1SO8  1U7T 	Sequence length (263 aa) 	---

CHISEL Cluster Query



CHISEL

high-resolution evolutionary analysis of enzymatic sequences

[Home](#)[Search](#)[Classify](#)[Contact us](#)[Download](#)[Help](#)

The classification page allows users to submit a sequence for prediction of function and possible taxonomy of the organism of origin

Due to large quantities of Chisel families, the classification will take several minutes. Please be patient. We are working on this to make the classification process quicker.

Upload the fasta file of sequences

Enter the sequences in fasta format

```
>Seq1
aftddrewsdweraallcdewqqqsdf
```

Search any sequence for
cluster similarity against the
HMM Chisel profiles

CHISEL Cluster Query Results

It takes a while for the results to be displayed

Query Sequence:	17432086		
Accession:	[none]		
Description:	PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE OXIDOREDUCTASE PROTEIN [Ralstonia solanacearum]		
<i>Model</i>	<i>Description</i>	<i>Score</i>	<i>E-value</i>
SF000147_21_B_Bacteria1	1.2.1.16 succinate-semialdehyde dehydrogenase [NAD(P)] (Bacterial version)	1231.5	0
SF000147_28_B_Bacteria1	1.2.1.16 succinate-semialdehyde dehydrogenase [NAD(P)] (Bacterial version)	1029.2	0
SF000147_8_B_Bacteria1	1.2.1.16 succinate-semialdehyde dehydrogenase [NAD(P)] (Bacterial version)	877.7	2.8e-261
SF000147_30_B_Rhizobiales1	1.2.1.16 succinate-semialdehyde dehydrogenase [NAD(P)] (Bacterial version)	823.8	4.7e-245
SF000147_29_E_Eukaryota1	1.2.1.24 succinate-semialdehyde dehydrogenase (Eukaryotic version)	709.8	9.5e-211
SF000147_17_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	447.4	9.6e-132
SF000147_1_B_Bacteria1	1.2.1.8 betaine-aldehyde dehydrogenase (Bacterial version)	432.6	2.7e-127
SF000147_23_B_Proteobacteria1	1.2.1.39 phenylacetaldehyde dehydrogenase (Bacterial version)	409.5	2.4e-120
SF000147_6_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	409.2	3e-120
SF000147_13_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	383.3	1.9e-112
SF000147_5_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	355.5	4.4e-104
SF000147_15_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	344.4	9.9e-101
SF000147_20_B_Proteobacteria1	1.2.1.22 lactaldehyde dehydrogenase (Bacterial version)	343.6	1.7e-100
SF000147_18_B_Bacteria1	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	315.6	4.4e-92
SF000147_17_B_Bacteria2	1.2.1.3 aldehyde dehydrogenase (NAD) (Bacterial version)	301.6	7.5e-88
SF000147_15_E_Eukaryota1	1.2.1.3 aldehyde dehydrogenase (NAD) (Eukaryotic version)	298.9	4.8e-87
SF000147_22_E_Eukaryota1	1.2.1.8 betaine-aldehyde dehydrogenase (Eukaryotic version)	285.1	6.8e-83
SF000147_4_B_Bacteria1	1.2.1.16 succinate-semialdehyde dehydrogenase [NAD(P)] (Bacterial version)	275.3	6.1e-80